Original Publication

🔓 Open Access

Navigate the Challenges of Simulation for Assessment: A Faculty Development Workshop

Megan A. Koster, MD, PharmD*, Morgan Soffler, MD

*Corresponding author: lovullo.m@gmail.com

Abstract

Introduction: Given barriers to learner assessment in the authentic clinical environment, simulated patient encounters are gaining attention as a valuable opportunity for competency assessment across the health professions. Simulation-based assessments offer advantages over traditional methods by providing realistic clinical scenarios through which a range of technical, analytical, and communication skills can be demonstrated. However, simulation for the purpose of assessment represents a paradigm shift with unique challenges, including preservation of a safe learning environment, standardization across learners, and application of valid assessment tools. Our goal was to create an interactive workshop to equip educators with the knowledge and skills needed to conduct assessments in a simulated environment. **Methods:** Participants engaged in a 90-minute workshop with large-group facilitated discussions and small-group activities for practical skill development. Facilitators guided attendees through a simulated grading exercise followed by in-depth analysis of three types of assessment tools. Participants designed a comprehensive simulation-based assessment encounter, including selection or creation of an assessment tool. **Results:** We have led two iterations of this workshop, including an in-person format at an international conference and a virtual format at our institution during the COVID-19 pandemic, with a total of 93 participants. Survey responses indicated strong overall ratings and impactfulness of the workshop. **Discussion:** Our workshop provides a practical, evidence-based framework to guide educators in the development of a simulation-based assessment program, including optimization of the environment, design of the simulated case, and utilization of meaningful, valid assessment tools.

Keywords
Simulation, Simulation Training, Assessment, Competency-Based Assessment, Interprofessional Education, Faculty Development

Educational Objectives

By the end of this activity, participants will be able to:

1. Evaluate a learner in a simulated patient encounter using an assessment instrument.
2. Evaluate the strengths and weaknesses of three types of assessment tools.
3. Design an aspect of a simulation-based assessment encounter, including creation of an assessment tool.
4. List best practices in optimizing simulation for the purpose of assessment, including the prebrief, simulated case, and debrief.

Introduction

Meaningful, feasible, and valid assessment of learners remains a challenge to educators across the health professions. Assessment is increasingly devoted to the identification of competence, which has been defined as the habitual use of communication, knowledge, technical, and clinical reasoning skills in daily practice and further codified, particularly in graduate medical education, with core competencies and associated developmental milestones.[1-3] However, an ideal means of assessing trainees for competencies has not yet been established.

Competence is dependent on context, clinical subject matter, and the developmental level of the learner.[4] Given this variability, competence may not be adequately assessed on a single encounter or by a single method. The literature suggests that the various methods of assessment—written examinations, bedside observations, or simulated encounters (standardized patients, objective structured clinical evaluations)—each have strengths

and weaknesses that can be leveraged to achieve specific goals.[4,5] The framework for clinical competence introduced by Miller provides a pyramidal structure for organizing assessment methods that is analogous to Bloom's taxonomy for learning.[6] The highest levels of Miller's pyramid are reserved for techniques that assess the most complex skills, specifically, demonstration of behavioral, analytical, and communication skills in simulated or live clinical environments.[5,6]

How can the individual educator or program administrator apply Miller's framework? One potential solution is to leverage existing resources at hospital systems or health professions schools for simulation-based assessment. Simulation has become embedded in health care professional education. In a 2011 survey of AAMC-affiliated teaching hospitals and US medical schools, simulation was used by nearly all responding institutions for the purpose of core competency education and often, though to a lesser degree, for the purpose of competency assessment.[7] Most reported simulation-based assessments were for learner feedback, rather than evaluation, and only rarely for certification—thus, for low-stakes assessments. In the decade since, new studies have emerged regarding the reliability, validity, and feasibility of simulation-based assessment tools, including checklists, global rating scales, and objective measurements such as time or number of actions performed.[8-10] However, simulation-based assessment has some intrinsic challenges, including cost, faculty training, preservation of a safe learning environment, and standardization across examinees.

The goal of our 90-minute workshop is to bridge the gap between the growing interest in simulation for assessment and the limited availability of resources to guide educators on practical development. Our workshop provides an interactive road map to effectively create a simulation-based assessment and navigate the associated challenges. The target audience includes educators across the health professions with an interest in simulation. There are simulation-based faculty development workshops published in *MedEdPORTAL* that serve as an introduction to simulation scenario design[11] and interprofessional simulation.[12] Our work expands upon these offerings by providing a framework to teach educators how to conduct assessments in a simulated environment. There are no other *MedEdPORTAL* publications that focus on faculty development regarding simulation for assessment.

Covered topics include recommendations for prebriefing and scenario design, as well as assessment tool validity, feasibility, and reliability. The session provides tips to optimize the environment for assessment, in-depth analysis of benefits

and trade-offs of various assessment tools, and an exercise to create a novel simulation-based assessment encounter. Small-group exercises and facilitated discussions are used to maximize participants' engagement and draw upon their prior experiences. We anticipate that by completion of this workshop, participants will emerge with strategies to design a simulation-based assessment encounter, including the simulation itself and the selection or creation of an appropriate assessment tool.

Methods

A proposal to develop this content was submitted for peer review by the Society for Simulation in Healthcare and selected for presentation to an in-person audience of interprofessional simulation educators at the International Meeting for Simulation in Healthcare (IMSH) 2020 conference. The workshop was subsequently presented virtually to a group of internal medicine physician educators of the Beth Israel Deaconess Medical Center in 2020. Prerequisite knowledge requirements for audience members included familiarity with simulation for health professional education.

As facilitators, we leveraged our collective experience as leaders in medical education and simulation at an academic medical center. The workshop authors were members of the Simulation Core Faculty at Beth Israel Deaconess Medical Center. Two out of three facilitators completed formal training at the Center for Medical Simulation and held leadership positions in simulation at the Carl J. Shapiro Center for Education at Beth Israel Deaconess Medical Center. Future iterations of this workshop could be offered by educators with similar training or leadership roles in simulation.

Setup

Our workshop was delivered via in-person and virtual formats (the latter was chosen due to the COVID-19 pandemic). For the in-person format, participants were seated at round or rectangular tables to facilitate small-group conversation. We had three facilitators and over 80 participants, with eight to 10 people per table. Projectors displayed slides (Appendix A) to anchor the content and provide directions, at a ratio of one projector per ∼40 persons. Facilitators had wireless microphones, and one handheld microphone was used by participants. Paper copies of the materials needed for activities were placed upon the tables. A different group sat at each table and received a copy of one assessment tool and one discussion guide:

- Pair 1: checklist tool and group 1 discussion guide.
- Pair 2: global rating scale tool and group 2 discussion guide.

- Pair 3: objective tool and group 3 discussion guide.

Half of the seats at each table received a blue index card. Participants at seats with a blue index card were instructed to stand up and switch tables at the midpoint of the session. This was done to mix up the groups for the second activity and to provide a stretch break mid-session. This technique could be replaced with any method to mix the groups for the second activity.

When delivered in virtual format, a shared online meeting platform was used that enabled breakout rooms, screen-sharing, and participant chat functions. Materials and breakout room assignments were distributed in advance of the session via email. We had two facilitators for ∼10 participants in the virtual session.

Content

Introduction: Appendix A includes slides used during the session, and Appendix B includes a workshop overview. Annotated slides with a facilitator script and other instructions can be found in Appendix C. The workshop began with facilitator introductions and disclosures and an opportunity to welcome the participants. Participants were encouraged to introduce themselves, their field, and their role in simulation to their tablemates. As a large group, participants were invited to raise their hands in response to a series of yes/no questions to provide attendees and facilitators with a sense of the demographics of the audience. Inevitably, only a few raised hands remained for participants who had previously conducted simulation-based assessments, and those participants were invited to briefly describe their experience. Next, participants were invited to debate pros and cons of simulation for the purpose of assessment based on their experiences or impressions. Facilitators were prepared to highlight commonly cited advantages and challenges and how these would be addressed over the course of the workshop.

Activity 1: In this activity, participants used an assessment tool to evaluate a learner depicted in a video of a fictitious assessment encounter. Each table had one of three assessment tools (Appendix D) that we created specifically for this activity. The goal was to demonstrate the strengths and weakness of the three most common types of assessment tools—checklist, global rating scale, and objective measurement tool—by having groups compare their experiences using the tools. We began with instructions for participants and explained the time allotted to complete each step. Then, we oriented participants to the circumstances shown in the video, including the level of the depicted learner, the clinical scenario, and the context/stakes

of this assessment. The video, embedded in the slides and also available as Appendix E, was played only once. This was done to demonstrate the challenges of live assessment and the necessity of rater training to improve accuracy of the tool. After watching the video, participants at each table compared their scores and examined aspects of practicality, content, and level of detail of the tool. Then, facilitators displayed each tool for the large group and invited tables that had used the tool to share their consensus on ease of use, interrater reliability, and whether the tool meaningfully captured the learner's performance. By design, the tools used in this exercise had imperfections; their flaws served to highlight the limitations of their format based on our experience and as reported in the literature.[4,5] Similarly, the actor in the video portrayed a learner with clear deficits in communication and professionalism to underscore how these aspects of performance may not be well captured with objective or checklist-style tools. At the conclusion of this activity, participants were given a 5-minute break, and those with a blue index card at their seat were instructed to switch to new tables to redistribute the groups.

Activity 2: The goal of this activity was to design an encounter for the purpose of assessment from start to finish. Each table was assigned one of three key elements: choosing what and how to assess (including the assessment tool), crafting the prebrief, and planning the simulation case itself. We presented the large group with a scenario of an interdisciplinary education and assessment need. Participants were asked to assume the role of a simulation education director for a large hospital recently assigned responsibility for designing a simulation-based program capable of training and assessing competent domestic violence screening for all patient-facing health care personnel. This scenario was chosen because it represented a skill (communication) demonstrable via simulation and relevant to an interprofessional audience. Paper discussion guides with the scenario and specific prompts for each group to consider were available on each table (Appendix F). Participants were given 15 minutes to work through their guide, and then, each topic was discussed as a large group. Facilitators highlighted principles of simulation and assessment relevant to each topic as supported by the literature. For example, for "choosing what and how to assess," we discussed selecting an observable standard of competence, ensuring that the means of evaluation had intact functional task alignment[13] and skill transferability, and choosing a tool with the appropriate degree of validity evidence for the stakes of the assessment.[14,15] For the prebrief, we emphasized the importance of always clearly disclosing the purpose and ground rules of the encounter, particularly if the simulation space

was used for both education and assessment. Lastly, regarding design of the simulation case itself, we highlighted the balance between signal and noise for the level of the learner, techniques of standardization to replicate the assessment across examinees, and the value of pilot testing.

Wrap-up: We summarized the workshop with discrete take-home pearls for assessment in a simulated environment and let the participants to ask questions. We provided our contact information and offered to review specific attendee needs at their home institutions. At the conclusion of the session, participants were given an online evaluation survey to evaluate their experience during the workshop. A sample evaluation form is shown in Appendix G.

Virtual Format
The previously described slides and activities were also used for the virtual format of our workshop. After the large-group introduction, disclosures, and pro/con debate, participants were advised to open the electronic version of the materials for activity 1, which had been sent in advance via email. Facilitators reviewed instructions for activity 1 and oriented the participants to their assessment tool and the circumstances of the video. As a large group, we watched the video while participants rated the learner with their assigned assessment tool. Then, participants were sorted into one of three breakout rooms for 10 minutes to compare experiences using their tool. The online meeting software allowed facilitators to enter each breakout room to check on progress and answer questions. Facilitators broadcast updates on the time remaining to all rooms prior to their closure. Debrief of each tool occurred via large-group discussion. We transitioned to the second activity after a brief stretch break. Again, all instructions and materials were reviewed as a large group. Participants were manually assigned into different breakout groups for the second activity. This was done to ensure that there was one person who worked with each assessment tool from activity 1 in each small group for activity 2 and to benefit from unique perspectives and experiences shaped by the first exercise. The postactivity debrief and wrap-up sections were identical to those in the in-person format. At the conclusion of the workshop, participants were offered the same online survey administered to the in-person attendees, via QR code.

Results

A total of 86 participants attended the in-person version of this workshop. Twenty-six participants (30%) completed the postworkshop survey, which was designed, conducted, and analyzed by the IMSH 2020 conference. The survey consisted of six items, each evaluated on a 5-point Likert scale. Scores were presented as a mean. Standard of error, standard deviations, and raw data were not made available to workshop authors. The results for all items were over 4 out of 5 (Table 1). The overall rating for the course was 4.7, where 1 = *poor* and 5 = *excellent.* Respondents rated the overall effectiveness of course faculty as 4.6, where 1 = *not at all effective* and 5 = *extremely effective.* In addition, respondents rated the workshop as highly impactful, 4.3 (1 = *not at all impactful,* 5 = *extremely impactful*), for application to their practice and for their team. Five participants (6%) left open-ended comments (Table 2). Three of these comments positively regarded the format, facilitation, and group exercises in the workshop. One comment was neutral; the participant indicated confusion about the intent of the workshop to address program evaluation versus individual learner assessment. One comment was negative; the participant felt the course did not help them to develop a better assessment tool. This comment may reflect that compared to other topics, the collection of validity evidence for assessment tools was not covered in depth.

A total of seven participants attended the virtual version of this workshop. All seven participants (100%) completed the postworkshop survey, which was conducted and analyzed by the workshop facilitators. The survey replicated the IMSH survey in format, questions, and answer choices with two exceptions. First, a different online survey platform (Qualtrics) was used. Second, one of the six items was eliminated as it was redundant for this population. The same 5-point Likert scales were used. The results for all items were over 4.6 out of 5 (Table 3). Overall ratings for the workshop, adherence to stated learning objectives, and overall effectiveness of workshop facilitators were 5 out of 5. Four participants (57%) provided open-ended comments (Table 4). Three comments praised the learning environment, quality of materials, and authentic group exercises. One comment provided feedback regarding technical aspects of launching the breakout rooms, suggesting randomly assigning rather than preassigning rooms for ease of facilitation.

Table 1. In-Person Workshop Survey Item Average Scores (*n* = 26)

Survey Item	Score
Overall rating for this course[a]	4.7
Adherence to stated learning objectives[a]	4.8
Did faculty verbally state their disclosures?[b]	4.8
How impactful will this content be for my practice?[c]	4.3
How impactful will this content be for my team?[c]	4.3
Overall effectiveness of course faculty[d]	4.6

[a]Rated on a 5-point Likert scale (1 = *poor,* 5 = *excellent*).
[b]Rated on a 5-point Likert scale (1 = *no,* 5 = *yes*).
[c]Rated on a 5-point Likert scale (1 = *not at all impactful,* 5 = *extremely impactful*).
[d]Rated on a 5-point Likert scale (1 = *not at all effective,* 5 = *extremely effective*).

Table 2. In-Person Workshop Survey Comments (*n* = 5)

Category	Comment
Positive	"Excellent with ideal lead-in and subsequent group work."
	"Great format."
	"This course was very helpful. Conversation was facilitated very well."
Neutral	"I thought it was about program assessment instead of learner assessment."
Negative	"Did not help me develop a better tool."

Discussion

To address the lack of practical resources on simulation-based assessment, we developed a comprehensive workshop to train participants on crucial skills needed to conduct assessments in the simulated environment. These skills included how to select and deploy an assessment tool to evaluate an individual learner, how to evaluate the appropriateness of a particular tool for the competency of interest, and how to optimize the prebrief and simulated case for the encounter.

When we embarked on designing this workshop, we were struck by the limited examples of simulation for the purpose of competency assessment in the literature.[8,9] Hart and colleagues provided the most comprehensive recent data, which included multicenter validity evidence for checklist and global rating scale tools designed to assess competencies preselected via modified Delphi technique.[9] This work demonstrates rigorous methodology appropriate for high-stakes simulation-based assessments. Keeping in mind that survey data suggested simulation is most commonly used for lower-stakes assessment (i.e., learner feedback or as a part of evaluation) but rarely for certification, we set out to build our workshop to be accessible and practical for these purposes.[7] This may be a limitation of our workshop—it has been designed to guide simulation-based educators in how to successfully optimize their practices from education for assessment but does not focus explicitly on collecting validity evidence for high-stakes certifications. Still, the workshop makes a foray into uncharted territory for educators and addresses a gap in available resources.

Table 3. Virtual Workshop Survey Item Average Scores (*n* = 7)

Survey Item	Score
Overall rating for this course[a]	5.0
Adherence to stated learning objectives[a]	5.0
Did faculty verbally state their disclosures?[b]	5.0
How impactful will this content be for my practice?[c]	4.6
Overall effectiveness of course faculty[d]	5.0

[a]Rated on a 5-point Likert scale (1 = *poor*, 5 = *excellent*).
[b]Rated on a 5-point Likert scale (1 = *no*, 5 = *yes*).
[c]Rated on a 5-point Likert scale (1 = *not at all impactful*, 5 = *extremely impactful*).
[d]Rated on a 5-point Likert scale (1 = *not at all effective*, 5 = *extremely effective*).

Table 4. Virtual Workshop Survey Item Results (*n* = 4)

Category	Comment
Positive	"The first breakout session on rating simulation was especially well done—really everyone came to understanding why hybrid evaluation would be best and the pros/cons of tools. Time management became an issue but only because of too much engagement."
	"Amazing preparation with small groups & materials ahead of time; very safe learning environment; truly stellar all around."
	"The materials were so well done. Everything felt very authentic to real challenges."
Constructive	"Consider random group assignments for logistics (would involve giving all prep materials to everyone and directing them to the correct one by room)."

In putting the workshop together, we recognized that an activity-based format was crucial to maintain participant engagement. We strove to demonstrate rather than lecture on our key take-home messages. The interactive elements increased in complexity throughout the workshop, beginning with reflecting on participant experiences in the pro/con debate, advancing to application with the assessment tool activity, and then culminating in designing an entire assessment encounter. In order to tackle designing an encounter, participants needed a working understanding of the strengths and limitations of various tools, so we intentionally arranged the tool activity first. This structure was well received by participants, as reflected in the positive survey comments from the in-person and virtual versions of the workshop.

In the design stages, we also paid specific attention to the multidisciplinary nature of our audience. For activity 1, participants evaluated a video portraying a resident physician as part of a fictitious annual performance assessment. Before the video played, we oriented the room to the context and the relevant details of the clinical scenario, such that participation in the activity did not require advanced medical knowledge. For activity 2, we intentionally chose a behavioral skill relevant to interdisciplinary practice. This proved beneficial, as our in-person audience included participants from medicine (both civilian and military), nursing, physical therapy, and respiratory therapy, and augments the generalizability of our workshop for *MedEdPORTAL*.

In addition, we highlight that the structure of our workshop, as proposed, exposes participants to different experiences in small groups. We rely on large-group discussion at the conclusion of each activity to create a cohesive set of learning points for all participants in order to augment their firsthand experiences. This format was chosen to maximize learning potential for the

whole group within the 90-minute time frame. However, we acknowledge that this design has limitations. Some small groups have more challenging activities than others, and the large-group debrief may not be enough to ensure a shared skill set. This disparity is particularly evident in activity 2, in which each small group designs an aspect of a simulation assessment encounter and one group has the more challenging task of designing two key elements of an assessment tool. Future facilitators can eliminate this imbalance by selecting an optional amendment to our format that requires 30 additional minutes of session time. In the optional amendment, detailed in Appendices B and C, small groups work through discussion guides 1-3 in sequence during 15-minute intervals, which are punctuated by 5-minute large-group debriefs. Activity 2 would then conclude with a 5-minute summative large-group discussion to review key points. If this modification is selected, participants would achieve a revised objective 3: Design an aspect of a simulation-based assessment encounter, including creation of an assessment tool.

Our work has other limitations. First, some of our recommendations for optimizing the environment for simulation-based assessment are based on our experiences. For example, our tips to announce the intent to evaluate in the prebrief and to leverage routine pilot-testing and training scripts are primarily derived from our observations. However, this approach is also an advantage, as it documents practical advice not available elsewhere. Second, raw data from the postworkshop survey were not made available to facilitators, so we do not have access to statistics that would be helpful to properly contextualize the results. In addition, the survey response rate following the in-person version of the workshop was low at 30%. Potential reasons for the low rate include lack of dedicated time to complete the survey in the session, lack of postworkshop reminder notifications to complete the survey, and survey fatigue among conference attendees. What we can take away from the results is that, on average, respondents found the facilitation effective, the content impactful, and the overall workshop excellent. Though the number of participants was small, these findings were similar between in-person and virtual formats.

In the future, this work could be expanded upon and informed by emerging literature on simulation-based assessments. As more examples of low- and high-stakes assessments enter the literature, practical features of well-designed assessments will crystalize. In the interim, our workshop can provide an accessible guide for interdisciplinary simulation educators to leverage their resources for competency assessment.

Appendices

- A. Navigate Challenges Slides.pptx
- B. Workshop Overview.docx
- C. Facilitator Guide.docx
- D. Activity 1 Assessment Tools.docx
- E. Activity 1 Video.mp4
- F. Activity 2 Discussion Guides.docx
- G. Evaluation Form.docx

All appendices are peer reviewed as integral parts of the Original Publication.

Megan A. Koster, MD, PharmD: Instructor in Medicine, Harvard Medical School and Division of Pulmonary and Critical Care Medicine, Mount Auburn Hospital

Morgan Soffler, MD: Instructor in Medicine, Harvard Medical School and Division of Pulmonary, Critical Care and Sleep Medicine, Beth Israel Deaconess Medical Center; Director of Simulation Research, Carl J. Shapiro Center for Education and Research, Beth Israel Deaconess Medical Center

References

1. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA.* 2002;287(2):226-235. https://doi.org/10.1001/jama.287.2.226

2. Batalden P, Leach D, Swing S, Dreyfus H, Dreyfus S. General competencies and accreditation in graduate medical education. *Health Aff (Millwood).* 2002;21(5):103-111. https://doi.org/10.1377/hlthaff.21.5.103

3. Edgar L, McLean S, Hogan SO, Hamstra S, Holmboe ES. *The Milestones Guidebook.* Accreditation Council for Graduate Medical Education; 2020. http://www.acgme.org/portals/0/milestonesguidebook.pdf

4. Epstein RM. Assessment in medical education. *N Engl J Med.* 2007;356(4):387-396. https://doi.org/10.1056/NEJMra054784

5. Wass V, Van der Vleuten C, Shatzer J, Jones R. Assessment of clinical competence. *Lancet.* 2001;357(9260):945-949. https://doi.org/10.1016/S0140-6736(00)04221-5

6. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med.* 1990;65(9):S63-S67. https://doi.org/10.1097/00001888-199009000-00045

7. Passiment M, Sacks H, Huang G. *Medical Simulation in Medical Education: Results of an AAMC Survey.* Association of American Medical Colleges; 2011. https://www.aamc.org/system/files/c/2/259760-medicalsimulationinmedicaleducationanaamcsurvey.pdf

8. Ryall T, Judd BK, Gordon CJ. Simulation-based assessments in health professional education: a systematic review. *J Multidiscip Healthc.* 2016;9:69-82. https://doi.org/10.2147/JMDH.S92695

9. Hart D, Bond W, Siegelman JN, et al. Simulation for assessment of milestones in emergency medicine residents. *Acad Emerg Med.* 2018;25(2):205-220. https://doi.org/10.1111/acem.13296

10. Murray DJ, Boulet JR, Avidan M, et al. Performance of residents and anesthesiologists in a simulation-based skill assessment. *Anesthesiology.* 2007;107(5):705-713. https://doi.org/10.1097/01.anes.0000286926.01083.9d

11. Robertson JM, Bradley D. Simulation clinical scenario design workshop for practicing clinicians. *MedEdPORTAL.* 2017;13:10645. https://doi.org/10.15766/mep_2374-8265.10645

12. Wong AH, Ruppel H, Gang M, Ng G. "The safety dance": a faculty development workshop partnering IPE and patient safety initiatives using simulation-based education. *MedEdPORTAL.* 2016;12:10440. https://doi.org/10.15766/mep_2374-8265.10440

13. Hamstra SJ, Brydges R, Hatala R, Zendejas B, Cook DA. Reconsidering fidelity in simulation-based training. *Acad Med.* 2014;89(3):387-392. https://doi.org/10.1097/ACM.0000000000000130

14. Cook DA, Hatala R. Validation of educational assessments: a primer for simulation and beyond. *Adv Simul (Lond).* 2016;1:31. https://doi.org/10.1186/s41077-016-0033-y

15. Cook DA, Brydges R, Ginsburg S, Hatala R. A contemporary approach to validity arguments: a practical guide to Kane's framework. *Med Educ.* 2015;49(6):560-575. https://doi.org/10.1111/medu.12678